

# PENGEMBANGAN *SOFTWARE* ANALISIS BUTIR SOAL YANG PRAKTIS DAN APLIKATIF

Ali Muhson, Barkah Lestari, Supriyanto, & Kiromim Baroroh

Universitas Negeri Yogyakarta, Jl. Colombo 1 Yogyakarta  
*e-mail*: alimuchson@yahoo.com

**Abstract: Developing a More Practical and Applicable Item-Analysis Software.** This study was aimed to develop item-analysis software, verify its feasibility, and identify the possible problems in using it, employing such research instruments as documentation, questionnaires, and interview guidelines to collect data. The result showed that the product was found to be more practical and applicable for classroom teachers, more specifically in the aspects of practicality and simplicity, usefulness, the substance, and the appearance. This study also revealed the challenges in the utilization of the product: Teachers appear to have little knowledge and skill in using *Microsoft Excel* program, analyzing items, and handling technical constraints.

**Keywords:** feasibility, item-analysis software, development

**Abstrak: Pengembangan *Software* Analisis Butir Soal yang Praktis dan Aplikatif.** Artikel hasil penelitian dan pengembangan ini memaparkan kelayakan program aplikasi Analisis Butir Soal (AnBuso). Data untuk uji kelayakan diperoleh melalui dokumentasi, kuesioner, dan wawancara, selanjutnya dianalisis secara deskriptif dan menjadi masukan untuk perbaikan. *Software* AnBuso dan buku panduan pengoperasian telah dimanfaatkan guru untuk melakukan analisis butir soal. *Software* dinilai layak dilihat dari aspek kepraktisan dan kemudahan, kebermanfaatan, substansi, dan tampilan. Kendala pemanfaatan *software* terkait dengan kurangnya penguasaan guru dalam pengoperasian program aplikasi *Microsoft Excel*, kurang terbiasa melakukan analisis butir, dan keterbatasan pemahaman konsep tentang analisis butir.

**Kata kunci:** AnBuso, kelayakan, analisis butir soal, pengembangan

Pendidikan merupakan kunci utama kemajuan bangsa, dan upaya peningkatan kualitas pendidikan harus terus dilakukan agar Indonesia mampu bersaing di kancah global. Potret kualitas pendidikan di Indonesia masih memprihatinkan. *Education Development Index (EDI)* Indonesia belum beranjak dari kategori sedang, yaitu berada pada peringkat ke-57 dari 115 negara (UNESCO, 2011), dan *Human Development Index (HDI)* Indonesia menempati peringkat ke-108 dari 187 negara (UNDP, 2014). Sementara itu, Singapura berada di posisi ke-9, Malaysia ke-62, dan Thailand ke-89. Hal ini menunjukkan bahwa kualitas pendidikan di Indonesia perlu ditingkatkan agar tidak tertinggal dari negara lain.

Proses pembelajaran menjadi bagian penting yang menentukan kualitas pendidikan secara keseluruhan, dan pendidik memiliki peran sentral dalam peningkatan kualitas tersebut. Peran pendidik tidak

hanya terbatas dalam merancang dan melaksanakan pembelajaran, melainkan juga dalam melakukan asesmen proses dan hasil belajar. Asesmen merupakan komponen penting dalam pembelajaran (Russell & Airasian, 2012), karena memiliki pengaruh kuat terhadap perbaikan proses pembelajaran (Raymond dkk., 2012; Bers, 2008). Penggunaan prosedur asesmen yang benar berkontribusi langsung terhadap peningkatan belajar peserta didik (Miller dkk., 2009). Untuk itu, guru perlu memiliki kemampuan mengembangkan instrumen asesmen yang mampu memotret secara tepat tingkat capaian kompetensi peserta didik.

Guna mengukur tingkat ketercapaian tujuan pendidikan, perlu dikembangkan instrumen asesmen yang mampu mengungkap seluruh komponen yang ingin dicapai dalam pembelajaran. Asesmen merupakan kegiatan pengumpulan bukti-bukti tentang pembelajaran siswa, dan berfungsi sebagai informasi untuk

pengambilan keputusan dalam pembelajaran (Stiggins & Chappius, 2012). Agar keputusan yang diambil tepat, asesmen harus memerhatikan keseluruhan aspek yang akan diukur agar mampu menggambarkan dengan tepat sasaran yang dituju.

Pemberlakuan kurikulum 2013 dalam implementasinya masih banyak mengalami permasalahan (Republika, 2014), misalnya sulitnya mengubah *mindset* guru (Metronews, 2014; Tempo, 2013). Hasil kajian menunjukkan bahwa 87 persen guru masih mengalami kesulitan dalam memahami cara melaksanakan asesmen autentik (Susilowati, 2013). Salah satu hal yang membuat guru “repot” adalah sistem asesmen yang memiliki terlalu banyak aspek (Tempo, 2014). Hal ini mengindikasikan bahwa kemampuan guru untuk melakukan asesmen secara baik masih perlu ditingkatkan.

Asesmen merupakan upaya sistematis untuk mengumpulkan, mengkaji, dan menggunakan informasi tentang program pendidikan dengan tujuan untuk meningkatkan dan mengembangkan pembelajaran (Banta dkk., 2014). Agar asesmen menghasilkan informasi yang tepat, perlu dilakukan pengumpulan bukti akurat terkait capaian hasil belajar siswa, dan menjadikan proses asesmen bermanfaat yaitu dapat meningkatkan motivasi dan prestasi belajar siswa (Stiggins & Chappuis, 2012). Informasi yang diperoleh melalui asesmen digunakan sebagai dasar untuk pengambilan keputusan tentang status siswa dalam kelompoknya dan menentukan langkah tindakan selanjutnya. Dalam melakukan asesmen pembelajaran perlu dirancang langkah-langkahnya secara rinci agar mampu memotret kompetensi siswa secara tepat.

Asesmen membantu guru memerjelas perumusan tujuan pembelajaran dan pencapaiannya, juga menciptakan pengalaman dalam menerapkan pengetahuan dengan konteks kehidupan nyata, dan memberikan berbagai cara bagi siswa untuk menunjukkan kemampuan dan keterampilannya (Darling-Hammond, 2014). Prosedur asesmen yang diterapkan secara benar dapat berkontribusi langsung terhadap peningkatan belajar siswa, yakni mengklarifikasi sifat hasil belajar, menyiapkan arah tujuan jangka pendek, memberikan balikan terhadap kemajuan belajar, memberikan informasi untuk mengatasi kesulitan belajar dan memilih pengalaman belajar masa depan, serta mengidentifikasi tujuan pembelajaran berikutnya (Miller dkk., 2009).

Asesmen bertujuan untuk menjaga keseimbangan kelas, digunakan untuk merencanakan dan melaksanakan pembelajaran, menempatkan siswa, memberikan balikan dan penghargaan, mendiagnosis permasalahan siswa, dan menilai tingkat kemajuan akademik (Russell & Airasian, 2012). Melalui asesmen dapat ditentukan rancangan pembelajaran berikutnya agar

prestasi akademik siswa dapat berkembang secara optimal.

Terdapat tiga tujuan utama dilakukannya asesmen (National Research Council, 1999). *Pertama*, asesmen sumatif dipergunakan untuk memantau kemajuan atau perbaikan pendidikan. Pendidik, pembuat kebijakan, orang tua, dan masyarakat ingin tahu berapa banyak siswa mencapai standar kinerja yang ditentukan. *Kedua*, asesmen formatif dipergunakan untuk memberikan balikan kepada guru dan siswa. Para guru menggunakan balikan untuk memperbaiki proses pembelajaran, dan siswa dapat menggunakan balikan untuk memantau capaian pembelajaran. *Ketiga*, asesmen akuntabilitas dipergunakan untuk mendorong perubahan dalam praktik dan kebijakan untuk mencapai perubahan yang diinginkan.

Asesmen kelas harus didasarkan pada beberapa kriteria, yaitu validitas, reliabilitas, fokus pada kompetensi, komprehensif, objektif, dan mendidik (Puskur, 2008; Anderson, 2003; Kubiszyn & Borich, 2013). Validitas dan reliabilitas menjadi bagian penting dalam kegiatan asesmen agar informasi yang dihasilkan sesuai dengan yang diharapkan.

Analisis butir menjadi bagian penting dalam menjamin validitas butir soal. *Item analyses play a somewhat more important role in construct and predictive validation* (Nunnally & Bernstein, 1994). Tiga hal yang diperhatikan dalam melakukan analisis butir adalah tingkat kesukaran, daya beda, dan distraktor. Tingkat kesukaran butir soal merupakan salah satu parameter yang sangat berguna, karena dapat mengetahui tingkat kualitas butir soal. Jika tingkat kesukaran mendekati 0, maka butir soal terlalu sukar, sedangkan jika tingkat kesukaran mendekati 1, maka soal terlalu mudah. Soal yang terlalu sukar atau terlalu mudah harus dihindari karena butir tersebut tidak dapat membedakan kemampuan siswa yang satu dengan siswa lainnya. Indeks kesukaran butir soal yang baik terletak dalam kategori sedang, yaitu pada interval 0,30-0,70 (Allen & Yen, 1979; Kaplan & Saccuzzo, 2005; Sudjana, 2011). Pada interval ini, informasi tentang kemampuan siswa akan diperoleh secara maksimal.

Kriteria kedua yang perlu diperhatikan adalah daya beda butir soal. Daya beda butir merujuk kepada kemampuan butir untuk membedakan peserta tes yang berkemampuan tinggi dan berkemampuan rendah. Untuk menentukan daya pembeda dapat dipergunakan indeks diskriminasi, indeks korelasi biserial, indeks korelasi *point biserial*, ataupun indeks korelasi. Indeks daya pembeda butir soal yang kecil nilainya menyebabkan butir tersebut tidak dapat membedakan antara siswa yang kemampuannya tinggi dan siswa yang kemampuannya rendah. Nilai daya

beda rendah menunjukkan adanya kemencengan distribusi skor dari populasi sebagai konsekuensi dari validitas tes yang rendah. Indeks daya beda dikatakan baik jika lebih besar atau sama dengan 0,3 (Nunnally & Bernstein, 2009; Kaplan & Saccuzzo, 2005; Azwar, 2003). Sementara itu koefisien daya beda antara 0,20-0,29 dianggap cukup baik (Alagumalai & Curtis, 2005), dan koefisien di bawah 0,2 dianggap tidak baik sehingga butir soal perlu dibuang (Ebel & Frisbie, 1991; Crocker & Algina, 2006).

Tes objektif bentuk pilihan ganda (*multiple choice*) dilengkapi dengan beberapa alternatif jawaban atau yang dikenal dengan istilah *option*. Jumlah *option* atau alternatif pilihan berkisar antara 3 sampai dengan 5 buah, dan dari kemungkinan-kemungkinan jawaban yang terpasang pada setiap butir, salah satu di antaranya merupakan jawaban benar (kunci jawaban) dan yang lainnya adalah jawaban salah yang dikenal dengan istilah pengecoh (*distractor*). Pada kenyataannya, dapat terjadi, dari alternatif yang terdapat pada butir tertentu sama sekali tidak dipilih oleh peserta tes. Hal demikian berarti alternatif tersebut tidak mampu berfungsi sebagai pengecoh yang baik. Pengecoh dinyatakan telah dapat berfungsi baik apabila distraktor tersebut sekurang-kurangnya dipilih oleh 5% dari seluruh peserta tes.

Asesmen akan bermakna dalam proses pembelajaran bilamana hasil asesmen dimanfaatkan dan ditindaklanjuti. Balikan asesmen merupakan bagian yang tidak terpisahkan dari suatu proses pembelajaran. Balikan bukanlah hal yang asing dalam dunia pendidikan (Irons, 2008), dan balikan akan sangat bermakna jika dilakukan secara tepat (Brookhart, 2008), karena dapat meningkatkan proses pembelajaran (Irons, 2008). Hasil penelitian menunjukkan bahwa balikan berdampak positif terhadap hasil belajar (James & Folorunso, 2012). Diperlukan kemampuan dan strategi khusus dalam melaksanakan asesmen yang baik serta memberikan balikan yang tepat agar mampu meningkatkan kualitas pembelajaran. Sudah selayaknya guru memiliki kompetensi yang memadai dalam melakukan asesmen dan balikan.

Balikan merupakan bukti yang menegaskan kebenaran suatu tindakan (Wiggins, 1993). Balikan dikonseptualisasikan sebagai informasi yang diberikan oleh perantara (misalnya, guru, teman sebaya, buku, orang tua, diri, atau pengalaman) mengenai aspek kinerja atau pemahaman seseorang (Hattie & Timperley, 2007). Balikan formatif dan sumatif adalah setiap informasi, proses, atau kegiatan yang memberi atau mempercepat belajar siswa berdasarkan komentar yang berkaitan dengan asesmen formatif dan asesmen sumatif (Irons, 2008). Balikan formatif disajikan sebagai informasi kepada peserta didik

dalam menanggapi beberapa tindakan peserta didik, misalnya berupa verifikasi akurasi respon, penjelasan jawaban yang benar, petunjuk, dan pemberian contoh (Shute, 2007).

Balikan merupakan sesuatu yang diberikan untuk mengkaji mengenai hal-hal yang telah dilakukan dan upaya untuk mengobservasi siswa berkaitan dengan bagaimana mereka melakukan aktivitasnya, serta apa yang harus dilakukan oleh guru untuk meningkatkan kemampuan mereka. Tujuan dari balikan formatif adalah untuk meningkatkan pengetahuan, keterampilan, dan pemahaman peserta didik, misalnya dalam hal pemecahan masalah (Shute, 2007), dan karenanya balikan harus bersifat interaktif, meningkatkan motivasi, dan terkait dengan upaya pemecahan masalah (Langer, 2011).

Siswa berpandangan bahwa balikan dapat memberikan informasi dan balikan yang baik kepada mereka (Peterson & Irving, 2008). Komentar yang diberikan dalam balikan formatif hanya dapat efektif bilamana siswa membaca dan memanfaatkannya (Higgins & Hartley, 2002). Selain itu, balikan dapat digunakan untuk meningkatkan keefektifan pembelajaran di kelas (Hattie & Timperley, 2007). Balikan yang diberikan kepada peserta didik secara tepat akan membantu mereka meningkatkan kinerjanya, memberikan ide tentang bagaimana mereka berkembang, meningkatkan motivasi, dan memberdayakan mereka sebagai peserta didik (Nair & Mertova, 2011). Dapat dikatakan bahwa balikan memiliki peran penting dalam proses pembelajaran, yaitu mengarahkan proses pembelajaran sesuai dengan tujuan yang ingin dicapai.

Asesmen pembelajaran merupakan kegiatan yang penting untuk dilakukan dan harus mampu mengukur secara tepat kompetensi peserta didik. Untuk itu, instrumen yang digunakan haruslah valid. Analisis butir soal memiliki peran penting untuk mengidentifikasi butir-butir soal yang baik. Asesmen selain memberikan informasi akurat mengenai kemampuan peserta didik, juga dapat mengidentifikasi lingkup materi mana yang dianggap sulit, dan lingkup materi yang belum dikuasai oleh peserta didik sehingga dapat dilakukan upaya perbaikan melalui kegiatan perbaikan.

Terdapat berbagai piranti lunak (*software*) analisis butir soal yang telah dikembangkan oleh para ahli namun belum dimanfaatkan secara optimal. Hal demikian terjadi karena sebagian besar *software* tersebut berbahasa asing sehingga dirasa sulit untuk memahami cara penggunaannya, cukup rumit, kurang praktis, dan kurang aplikatif. Selain itu, informasi yang diberikan pada berbagai *software* tersebut ditampilkan dalam format yang sangat beragam sehingga memersulit pengguna untuk menguasainya atau memahaminya. Untuk membantu guru dalam

kegiatan asesmen, penelitian ini (Muhson dkk., 2013) telah mengembangkan prototipe *software* Analisis Butir Soal (AnBuso) yang praktis dan aplikatif untuk menganalisis butir soal sesuai dengan kebutuhan guru di lapangan.

Melalui penggunaan *software* AnBuso dapat diketahui baik tidaknya soal yang dibuat oleh guru dari sisi daya beda, tingkat kesulitan, dan keefektifan distraktornya. Di samping itu, *software* juga memberikan informasi tentang kemampuan seluruh siswa dan tingkat ketercapaian Kriteria Ketuntasan Minimal (KKM). *Software* ini juga dirancang untuk mengidentifikasi dan mengelompokkan siswa yang termasuk ke dalam program remedial berdasarkan materi yang belum dikuasai sehingga akan mempermudah guru dalam pelaksanaan program remedial. Semua hasil analisis tersebut ditampilkan dan dapat dicetak dalam format laporan yang sangat mudah dibaca dan diinterpretasikan.

Hasil uji coba terbatas menunjukkan bahwa keberadaan *software* AnBuso disambut positif oleh guru untuk melakukan analisis butir soal yang mudah digunakan, praktis, dan aplikatif (Muhson, dkk, 2013). Sebagai produk awal, *software* tersebut perlu dikaji dan diujicobakan dengan melibatkan subjek penelitian baik guru dan pengawas untuk memperoleh masukan yang lebih komprehensif. Penelitian ini dimaksudkan untuk mengetahui kelayakan dan mengidentifikasi kendala yang dihadapi dalam penggunaan *software* AnBuso.

## METODE

Rancangan yang dipergunakan dalam penelitian ini adalah penelitian dan pengembangan dengan prosedur penelitian yang meliputi perancangan, pengembangan produk, validasi produk, uji coba produk, dan diseminasi produk. Kegiatan perancangan dan pengembangan produk telah dilakukan, dan artikel hasil penelitian ini memaparkan hasil uji coba produk pada khalayak yang lebih luas dan memperoleh informasi dan masukan yang komprehensif untuk penyempurnaan.

Penelitian dilakukan dengan melibatkan guru-guru dan pengawas sekolah di Daerah Istimewa Yogyakarta (DIY). Subjek penelitian sebanyak 65 orang yang berasal dari lima kabupaten/kota di Provinsi DIY. Subjek ditentukan secara purposif dengan mempertimbangkan kemampuan dalam pengoperasian komputer, khususnya program aplikasi *Microsoft Excel*.

Pengumpulan data dilakukan dengan teknik observasi, dokumentasi, angket, dan wawancara. Observasi digunakan untuk memperoleh data tentang kemam-

puan guru dalam menggunakan program aplikasi yang telah dikembangkan. Dokumentasi digunakan untuk memperoleh data yang berasal dari instrumen tes yang dilakukan oleh guru untuk mengukur kemampuan peserta didik, baik tes formatif maupun tes sumatif. Dokumen tersebut berupa soal-soal ujian dan ulangan harian, program remedial, dan hasil analisis butir soal yang selama ini digunakan guru.

Angket digunakan untuk mengidentifikasi masukan yang berasal dari guru, pengawas, pejabat Dinas Pendidikan, dan pakar. Angket juga digunakan untuk mengetahui tingkat kelayakan *software* yang telah dikembangkan mengenai aspek tampilan, substansi materi atau isi, kebermanfaatan, kepraktisan, dan kemudahan. Wawancara dilakukan kepada guru, pengawas, pejabat Dinas Pendidikan, dan pakar untuk memperoleh informasi mengenai kelebihan dan kelemahan dari *software* sebagai bahan penyempurnaan.

Data kemudian dianalisis secara deskriptif untuk menentukan kelayakan *software* dari aspek tampilan, substansi materi atau isi, kepraktisan, dan kemudahan. Kategori kelayakan *software* diklasifikasikan sebagaimana Tabel 1.

**Tabel 1. Kategorisasi Penilaian Kelayakan *Software***

No.	Skor	Kategori	Keterangan
1.	Lebih besar dari $M + 1,8 SD$	Sangat layak	$M = Mean$ (rata-rata)
2.	$M + 0,6 SD$ s.d. $M + 1,8 SD$	Layak	$SD = Standard Deviation$ (simpangan baku)
3.	$M - 0,6 SD$ s.d. $M + 0,6 SD$	Cukup	
4.	$M - 1,8 SD$ s.d. $M - 0,6 SD$	Tidak layak	
5.	Lebih kecil dari $M - 1,8 SD$	Sangat tidak layak	

## HASIL DAN PEMBAHASAN

Hasil uji coba pengembangan *software* melibatkan para pengguna seperti guru dan pengawas pendidikan sebanyak 65 orang, dengan rincian 72% sebagai pengajar di tingkat SLTA, 68% berasal dari sekolah negeri, dan 57% berstatus PNS. Subjek penelitian berasal dari berbagai guru mata pelajaran, yaitu Ekonomi, Akuntansi, Matematika, IPS, Bahasa Inggris, Bahasa Indonesia, Teknologi Informasi, Kimia, Fisika, Biologi, Geografi, dan Bahasa Arab.

Hasil penelitian menunjukkan bahwa kemauan guru untuk melakukan analisis butir soal masih belum sesuai harapan, yaitu sebanyak 11% selalu melakukan analisis butir soal, sebanyak 57% kadang-kadang melakukan analisis, dan sebanyak 12% tidak pernah melakukan analisis butir soal. Hal ini tentu penting

untuk dikaji lebih lanjut mengapa guru sebagai pendidik memiliki kemauan yang rendah dalam melakukan analisis butir soal. Umumnya, mereka hanya melakukan analisis butir soal jika dituntut oleh pengawas. Hal demikian terjadi karena kebanyakan guru kurang menguasai *software* analisis butir soal yang sudah ada. Kalaupun menguasai, tampaknya mereka tidak mampu membangkitkan minat dan kemauan untuk melakukan analisis butir soal secara berkelanjutan. Hal ini tentu berakibat pada tidak diketahuinya tingkat kualitas butir soal yang dikembangkan oleh guru.

Pengembangan *software* AnBuso dan buku panduannya telah diperbaiki berdasarkan masukan subjek penelitian. Buku panduan dikembangkan untuk memudahkan pengguna dalam memanfaatkan *software* ini. Buku panduan berisi informasi mengenai langkah-

langkah atau cara penggunaan *software* untuk memudahkan pengguna dalam pemanfaatannya. Panduan ini berisi pendahuluan, kerangka isi, data *input*, dan data laporan.

Beberapa perubahan penting yang dihasilkan adalah melakukan penyesuaian *software* dengan pemberlakuan Kurikulum 2013 yang terkait dengan aspek penilaian. Perbaikan yang dilakukan meliputi perubahan tampilan, *sheet* Input01, *sheet* Laporan Peserta, *sheet* Peserta Remedial, dan Perubahan Formula.

Berdasarkan masukan, perubahan tampilan perlu dilakukan karena dianggap terlalu banyak variasi warna sehingga terlihat kurang menarik. Perubahan tampilan tidak hanya dilakukan pada *sheet* Input (Gambar 1), namun juga dilakukan perubahan tampilan pada *sheet* Laporan (Gambar 2).

(a) Sebelum Perubahan

(b) Sesudah Perubahan

Gambar 1. Tampilan *Sheet* Input01 Sebelum dan Sesudah Perubahan

(a) Sebelum Perubahan

(b) Sesudah Perubahan

Gambar 2. Tampilan *Sheet* Laporan Butir Sebelum dan Sesudah Perubahan

AnBuso 5.1 For Teacher  
© 2011-2014 by Ali Muhson

(Hanya diperkenankan mengisi data atau menghapus tetapi tidak boleh memindah isi data atau menggunakan fasilitas Cut Paste)

Keterangan	Kolom Pengisian	VALIDASI
Tempat Laporan	Yogyakarta	OK
Tanggal Laporan	20 Agustus 2014	OK
Skala Penilaian (4, 10 atau 100)	4	OK
Nilai KKM	2.66	OK

**Jumlah dan Bobot Soal**

Jumlah soal pilihan ganda (Max 50)	20	OK
Jumlah soal essay (Max 10)	5	OK
Bobot soal pilihan ganda	60%	OK
Bobot soal essay	40%	OK

**Data Soal Pilihan Ganda**

Jumlah Alternatif Jawaban (Max 5)	4	OK
-----------------------------------	---	----

Input 01 Input 02 Lap Peserta Lap Butir Pola Jawaban Butir Lap Essay Materi Remedial Pese

Gambar 3. Penyesuaian Skala Penilaian

AnBuso 5.1 For Teacher  
© 2011-2014 by Ali Muhson

(Hanya diperkenankan mengisi data atau menghapus tetapi tidak boleh memindah isi data atau menggunakan fasilitas Cut Paste)

Keterangan	Kolom Pengisian	VALIDASI
Tempat Laporan	Yogyakarta	OK
Tanggal Laporan	20 Agustus 2014	OK
Skala Penilaian (4, 10 atau 100)	4	OK
Nilai KKM	2.66	OK

**Jumlah dan Bobot Soal**

Jumlah soal pilihan ganda (Max 50)	20	OK
Jumlah soal essay (Max 10)	5	OK
Bobot soal pilihan ganda	60%	OK
Bobot soal essay	40%	OK

**Data Soal Pilihan Ganda**

Jumlah Alternatif Jawaban (Max 5)	4	OK
-----------------------------------	---	----

Input 01 Input 02 Lap Peserta Lap Butir Pola Jawaban Butir Lap Essay Materi Remedial Pese

Gambar 4. Penambahan Bobot Penilaian

Sebagai konsekuensi pemberlakuan Kurikulum 2013, dilakukan penyesuaian pada *software* AnBuso untuk mengakomodasikan kepentingan guru dalam melaksanakan penilaian sesuai dengan Kurikulum 2013. Penyesuaian tersebut terkait dengan penentuan skala penilaian. Pada *software* versi awal, skala penilaian yang disediakan hanya 1-10 dan 1-100, sementara Kurikulum 2013 juga menggunakan skala penilaian 1-4. Hasil penyesuaian *software* dengan pilihan skala penilaian disajikan pada Gambar 3.

Hasil uji coba dengan subjek penelitian yang lebih besar menemukan kendala mengenai penentuan bobot penilaian antara soal objektif dan soal esai.

Perbaikan pada *software* dilakukan dengan menyediakan kolom untuk pembobotan, sebagaimana disajikan pada Gambar 4. Dengan perbaikan tersebut, diharapkan *software* akan semakin praktis dan aplikatif dalam penggunaannya. *Software* juga mungkin penggunaannya hanya untuk soal objektif atau soal esai saja, tidak harus kedua-duanya.

Penyesuaian terhadap Kurikulum 2013 berdampak pada perlunya perbaikan pada *sheet* Laporan Peserta. Pada bagian ini disajikan hasil perolehan nilai peserta tes sesuai dengan Kurikulum 2013 dilengkapi dengan predikatnya, sebagaimana disajikan pada Gambar 5. Bagian lain tidak mengalami perubahan karena telah sesuai dengan harapan guru.

DAFTAR NILAI SISWA									
Satuan Pendidikan		: SMA Negeri 80 Yogyakarta							
Nama Tes		: Sumatif							
Mata Pelajaran		: Ekonomi							
Kelas/Program		: XI/IPS							
Tanggal Tes		: 30 Agustus 2014							
Pokok Bahasan/Sub		:							
								<b>KKM</b>	
								2.66	
No	NAMA PESERTA	L/P	HASIL TES OBJEKTIF			NILAI TES ESSAY	NILAI AKHIR	PREDIKAT	KET
			BENAR	SALAH	NILAI				
1	RIANDITA PRATIWI	P	19	1	3.80	3.51	3.68	A	Tuntas
2	N VIAN BURHANUDIN L	L	13	7	2.60	2.77	2.67	B	Tuntas
3	DWI CAHYO ADHI	L	12	8	2.40	3.14	2.70	B	Tuntas
4	ISHMATUL FATHIYAH	P	12	8	2.40	3.32	2.77	B	Tuntas
5	BAGUS PRATOMO ARIF AL HAKIM	L	11	9	2.20	2.89	2.48	B-	Belum tuntas
6	R BRATA KUSUMASMARA	L	15	5	3.00	3.14	3.06	B+	Tuntas
7	LILIANA ROMANTIKA SARI	P	12	8	2.40	3.02	2.65	B-	Belum tuntas
8	RIYAN ABI SAYEKTI	P	18	2	3.60	2.52	3.17	B+	Tuntas
9	DWI TITIK PUSPITOSARI	P	12	8	2.40	2.65	2.50	B-	Belum tuntas
10	PIPIT HARYADI	L	11	9	2.20	2.58	2.35	B-	Belum tuntas
11	AWIG JUWANTO	L	17	3	3.40	2.65	3.10	B+	Tuntas
12	ELIN DAHRIAH	P	15	5	3.00	2.71	2.88	B	Tuntas

Gambar 5. Perbaikan Sheet Laporan Peserta

PENGELOMPOKAN PESERTA REMEDIAL							
Satuan Pendidikan		: SMA Negeri 80 Yogyakarta					
Nama Tes		: Sumatif					
Mata Pelajaran		: Ekonomi					
Kelas/Program		: XI/IPS					
Tanggal Tes		: 30 Agustus 2014					
Pokok Bahasan/Sub		:					
No	Kompetensi Dasar	Peserta Remedial	Hari	Tgl	Jam	Tempat	
1	Masalah Pokok Ekonomi	DWI CAHYO ADHI ; LILIANA ROMANTIKA SARI ; ELIN DAHRIAH; FITRIYANTI SYAH;					
2	Kelangkaan	Tidak Ada					
3	Sistem ekonomi	N VIAN BURHANUDIN L ; DWI CAHYO ADHI ; FITRIYANTI SYAH; TINA RETNANINGTYAS; NUR SABAR SETIANI;					
4	Perilaku Konsumen	N VIAN BURHANUDIN L ; DWI CAHYO ADHI ; ISHMATUL FATHIYAH; BAGUS PRATOMO ARIF AL HAKIM ; DWI TITIK PUSPITOSARI; FITRIYANTI SYAH; TINA RETNANINGTYAS; SITI SAONAH;					
5	Kebijakan Moneter	DWI CAHYO ADHI ; LILIANA ROMANTIKA SARI ; NING TYAS TRIASTUTI; NURLITA GADISA R;					
6	Elastisitas Permintaan	N VIAN BURHANUDIN L : DWI CAHYO ADHI : R BRATA					

Gambar 6. Perbaikan Sheet Laporan Peserta

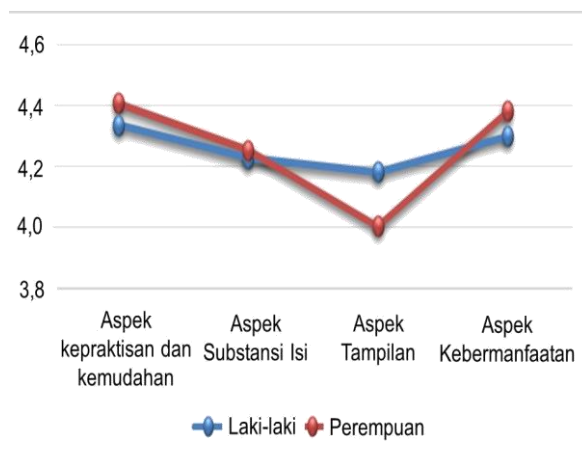
Tujuan utama dilakukannya analisis butir soal adalah untuk mengetahui kualitas butir soal. Selain itu, informasi hasil penilaian dapat dimanfaatkan untuk tindak lanjut tindakan baik berupa tindakan remedial maupun pengayaan. Hasil laporan penilaian juga harus mampu mengategorikan peserta yang perlu memperoleh pengayaan dan remedial. *Software* versi awal telah mengantisipasi hal tersebut namun hanya sebatas pengelompokan peserta pada tindakan remedial saja dan belum disediakan kolom untuk melakukan tindak lanjut. Untuk itu, perbaikan dilakukan,

tidak hanya menentukan kelompok peserta remedial atas dasar pengukuran capaian hasil pembelajaran, namun juga disediakan kolom untuk pengisian jadwal kegiatan remedialnya pada masing-masing kompetensi dasar. Hal itu untuk mengamodasikan kebutuhan kepraktisan dalam menentukan jadwal remedial. Perbaikan pada *sheet* Laporan Peserta Remedial disajikan pada Gambar 6.

Hasil uji kelayakan memperlihatkan bahwa *software* AnBuso yang dikembangkan dinilai sangat layak oleh sebanyak 46% subjek penelitian, 51% meny-

takan layak, dan 3% menyatakan cukup layak. *Software* yang dikembangkan dirasakan memiliki manfaat dan sangat membantu guru dalam melakukan analisis butir soal.

Nilai kelayakan tertinggi pada aspek kepraktisan dan kemudahan, dan aspek kebermanfaatan, serta nilai kelayakan terendah pada aspek tampilan sebagaimana disajikan pada Gambar 7. Selain itu, tidak terdapat perbedaan nilai kelayakan dilihat dari jenis kelamin subjek penelitian. *Software* AnBuso dianggap praktis, mudah digunakan, dan bermanfaat dalam membantu penyelesaian analisis butir soal dengan informasi hasil analisis yang lengkap.



**Gambar 7. Hasil Uji Kelayakan Software**

Aspek tampilan dinilai paling rendah dibandingkan dengan aspek lainnya. Hal demikian disebabkan keterbatasan dalam mengatur tampilan karena program aplikasi ini tidak berdiri sendiri melainkan melekat dan disesuaikan dengan fitur yang tersedia dalam *Microsoft Excel*. Komponen yang dinilai rendah adalah pemilihan warna, tata letak, dan pemilihan jenis huruf (*font*). Untuk itu, ke depan *software* ini masih perlu dilakukan perbaikan dari sisi tampilan. Pemilihan variasi warna dan *font* perlu dipertimbangkan agar diperoleh tampilan yang lebih menarik. Bahkan bila diperlukan dengan menggunakan program desain grafis dalam merancang tampilan.

Aspek substansi isi *software* dinilai sangat baik karena sesuai dengan kebutuhan. Selain itu, *software* juga dinilai praktis untuk digunakan, menarik, inovatif, kreatif, interaktif, dan unik. Informasi yang dihasilkan sangat lengkap, tidak hanya menampilkan hasil analisis butir soal objektif dan esai, juga menampilkan hasil capaian nilai dan KKM peserta didik. Hasil analisis juga menyajikan materi-materi tertentu yang belum dikuasai oleh masing-masing peserta didik untuk kepentingan tindakan remedial dan pengayaan.

Dari aspek kepraktisan dan kemudahan, *software* dinilai sangat baik. *Software* dinilai mudah digunakan, dipahami, dipelajari, dibaca, dan diinterpretasikan hasilnya. Untuk memanfaatkannya, cukup menggunakan pengetahuan dan keterampilan yang dimiliki mengenai program aplikasi *Microsoft Excel*. Program ini umumnya sudah dikuasai oleh guru, sehingga memudahkan dalam pemanfaatan dan aplikasinya.

*Software* dinilai memiliki manfaat yang tinggi karena aplikatif dan sesuai dengan kebutuhan guru. Keberadaan *software* membantu guru untuk melakukan analisis butir soal, dan tampilan hasil analisis sangat sesuai dengan kebutuhan karena sekaligus dapat melakukan evaluasi terhadap kualitas instrumen dan mengukur capaian kompetensi peserta didik. Selain itu, ketersediaan format laporan hasil analisis juga dapat dipergunakan untuk keperluan laporan administrasi guru.

Dalam kenyataannya, kendala yang dihadapi adalah masih ada beberapa guru yang kurang mahir dalam penguasaan komputer, khususnya penguasaan pada program aplikasi *Microsoft Excel* yang masih terbatas. Banyak menu dan fasilitas yang disediakan di *Microsoft Excel* namun belum dimanfaatkan secara optimal. Dikarenakan *software* ini terintegrasi dengan program aplikasi *Microsoft Excel*, akibatnya beberapa guru kurang lancar dalam pemanfaatannya.

Kebersediaan dan kebiasaan guru untuk melakukan analisis butir soal masih dalam kategori jarang. Hanya sedikit guru yang senantiasa melakukan analisis butir soal, baik terhadap soal harian, soal semesteran, maupun soal ujian akhir. Sebagai akibatnya, guru kurang lancar dalam melakukan analisis butir soal karena belum terbiasa. Selain itu, pengetahuan dan penguasaan guru tentang konsep analisis butir soal juga masih terbatas. Sementara pengembangan *software* didasarkan pada konsep analisis butir soal terstandar, akibatnya pemahaman guru terhadap angka-angka hasil analisis masih kurang. Walaupun hasil analisis sudah dikemas dalam bentuk laporan yang siap ditandatangani, masih perlu dijelaskan tentang makna dari hasil analisis tersebut.

## SIMPULAN

Penelitian telah berhasil mengembangkan *software* AnBuso yang dapat dimanfaatkan oleh guru untuk melakukan analisis butir soal secara praktis dan aplikatif. *Software* dibuat terintegrasi dengan program aplikasi *Microsoft Excel* yang di dalamnya terdapat *sheet* untuk *input* data, *sheet* pengolahan data, dan *sheet* laporan hasil analisis dalam bentuk tabel dan gambar.



*Software* yang dihasilkan dinilai sangat layak oleh guru dari aspek kepraktisan dan kemudahan, aspek kebermanfaatannya, aspek substansi isi, dan aspek tampilan. Kendala yang dihadapi oleh beberapa guru dalam penggunaan *software* ini adalah kurangnya penguasaan guru pada program aplikasi *Microsoft Excel*, kurang terbiasanya melakukan analisis butir soal, keterbatasan pemahaman konsep analisis butir soal, dan ketidاكلancaran pengoperasian komputer.

*Software* ini dinilai sangat layak dan sangat bermanfaat, oleh karena itu perlu diseminasi yang lebih luas tentang penggunaannya agar lebih dikenal, dan lebih lanjut mampu meningkatkan kinerja guru dalam melakukan analisis butir soal yang berdampak pada peningkatan capaian pembelajaran siswa. Pengembangan *software* masih perlu terus dilakukan agar mampu memenuhi perkembangan kebutuhan guru terkait dengan analisis butir soal.

## DAFTAR RUJUKAN

- Alagumalai, S. & Curtis, D.D. 2005. Classical Test Theory. Dalam Alagumalai, S., Curtis, D.D., & Hungi, N. (Eds.), *Applied Rasch Measurement: A Book of Exemplars* (hlm 1-13). Norwell, MA: Springer.
- Allen, M.J. & Yen, W.M. 1979. *Introduction to Measurement Theory*. Monterey, CA: Brooks/Cole Publishing Company.
- Anderson, L.W. 2003. *Classroom Assessment: Enhancing the Quality of Teacher Decision Making*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Azwar, S. 2003. *Sikap Manusia: Teori dan Pengukurannya* (Edisi ke-2). Yogyakarta: Pustaka Pelajar.
- Banta, T.W., Palomba, C.A., & Kinzie, J. 2014. *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*. San Fransisco: Jossey-Bass.
- Bers, T.H. 2008. The Role of Institutional Assessment in Assessing Student Learning Outcomes. *New Directions for Higher Education*, 141: 31-39.
- Brookhart, S.M. 2008. *How to Give Effective Feedback to Your Students*. Virginia: Association for Supervision and Curriculum Development.
- Crocker, L. & Algina, J. 2008. *Introduction to Classical and Modern Test Theory*. Ohio: Cengage Learning.
- Darling-Hammond, L. 2014. *Next Generation Assessment: Moving Beyond the Bubble Test to Support 21st Century Learning*. San Fransisco: Jossey-Bass.
- Ebel, R.L. & Frisbie, D.A. 1991. *Essentials of Educational Measurement*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Hattie J. & Timperley, H. 2007. The Power of Feedback. *Review of Educational Research*, 77 (1): 81-112.
- Higgins, R. & Hartley, P. 2002. The Conscientious Consumer: Reconsidering the Role of Assessment Feedback in Student Learning. *Studies in Higher Education*, 27 (1): 53-64.
- Irons, A. 2008. *Enhancing Learning through Formative Assessment and Feedback*. New York: Routledge.
- James, A.O. & Folorunso, A.M. 2012. Effect of Feedback and Remediation on Students' Achievement in Junior Secondary School Mathematics. *International Education Studies*, 5 (5): 153-162.
- Kaplan, R.M. & Saccuzzo, D.P. 2005. *Psychological Testing: Principles, Applications, and Issues* (6th edition). Belmont, CA: Thomson Wadsworth.
- Kubiszyn, T. & Borich, G.D. 2013. *Educational Testing and Measurement: Classroom Application and Practice* (10<sup>th</sup> edition). Hoboken, NJ: John Wiley & Sons, Inc.
- Langer, P. 2011. The Use of Feedback in Education: A Complex Instructional Strategy. *Psychological Reports*, 109 (3): 775-784.
- Metronews. 2014. *Ini Delapan Masalah dalam Implementasi Kurikulum 2013*, (Online), (www.Metronews.com), diakses 19 Oktober 2014.
- Miller, M.D., Linn, R.L., & Gronlund, N.E. 2009. *Measurement and Assessment in Teaching* (tenth edition). New Jersey: Pearson Education Inc.
- Muhson, A., Lestari, B., Supriyanto, & Baroroh, K. 2013. *Pengembangan Software AnBuso sebagai Solusi Alternatif bagi Guru dalam Melakukan Analisis Butir Soal Secara Praktis dan Aplikatif*. Laporan Penelitian tidak diterbitkan. Yogyakarta: LPPM UNY.
- Nair, C.S. & Mertova, P. 2011. *Student Feedback: The Cornerstone to an Effective Quality Assurance System in Higher Education*. New Delhi: Oxford Cambridge.
- National Research Council 1999. *The Assessment of Science Meets the Science of Assessment. Board on Testing and Assessment Commission on Behavioral and Social Sciences and Education, National Research Council*. Washington, DC: National Academy Press.
- Nunnally, J.C. & Bernstein, I.H. 1994. *Psychometric Theory* (Third Edition). New York: McGraw-Hill, Inc.
- Peterson, E.R. & Irving, S.E. 2008. Secondary School Students' Conceptions of Assessment and Feedback. *Learning and Instruction*, 18: 238-250.
- Puskur 2008. *Model Penilaian Kelas Kurikulum Berbasis Kompetensi*. Jakarta: Balitbang Depdiknas.
- Raymond, J.E., Homer, C.S.E., Smith, R. & Gray, J.E. 2012. Learning through Authentic Assessment: An Evaluation of a New Development in the Undergraduate Midwifery Curriculum. *Nurse Education in Practice*, 30: 1-6.
- Republika. 2014. *Implementasi Kurikulum 2013 Masih Dibayangi Banyak Masalah*, (Online), (www.republika.co.id), diakses 18 November 2013.

- Russell, M.K. & Airasian, P.W. 2012. *Classroom Assessment: Concepts and Applications* (7<sup>th</sup> edition). New York: McGraw-Hill.
- Shute, V.J. 2007. *Focus on Formative Feedback*. Unpublished Research Report. Princeton, NJ: Educational Testing Service (ETS).
- Stiggins, R.J. & Chappuis, J. 2012. *An Introduction to Student Involved Assessment for Learning* (Sixth edition). Boston: Pearson assessment training institute.
- Sudjana, N. 2011. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.
- Susilowati. 2013. *Kurikulum 2013, 87 Persen Guru Kesulitan Cara Penilaian*, (Online), (<http://unnes.ac.id>), diakses 18 November 2013.
- Tempo. 2013. *Problematika implementasi Kurikulum 2013*, (Online), ([www.tempo.co](http://www.tempo.co)), diakses 10 Juli 2013.
- Tempo. 2014. *Kurikulum 2013, Apa Saja Kendalanya?* (Online), ([www.tempo.co](http://www.tempo.co)), diakses 16 Agustus 2014.
- UNDP. 2014. *2014 Human Development Report*, (Online), (<http://www.undp.org/content/undp/en/home/librarypage/hdr/2014-human-development-report.html>), diakses 5 Maret 2014.
- UNESCO. 2011. *Education for All: Global Monitoring Report*, (Online), (<http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/efareport/statistics/efa-development-index/>), diakses 5 Maret 2014.
- Wiggins, G.P. 1993. *Assessing Student Performance: Exploring the Purpose and Limits of Testing*. San Francisco: Jossey-Bass Publishers.